MVA'94 IAPR Workshop on Machine Vision Applications Dec. 13-15, 1994, Kawasaki

An Interactive Map Drawing Recognition System with Learning Ability

Wei Lu, Masao Sakauchi
Institute of Industrial Science
University of Tokyo, Japan

7-22-1, Roppongi, Minato-ku,
Tokyo, 106 Japan

Abstract

We have developed an interactive map drawing recognition system which interactively process graphical objects that are difficult for fully automated recognition. It generates recognition proposals to reduce operation efforts. In this paper, we propose a new interactive recognition system which learns from examples processed during interactive operation so as to further improve the success rate of proposals. The system also plans the way the examples are presented to operator for judgement to realize fast and stable convergence of learning. The planning algorithm is also applicable to training of neural networks.

Experimental results show that as a result of the proposed algorithms, the efficiency of the system is further improved when compared to the existing one proposed by authors group before (success rate of proposal from about 70% to greater than 90%); the convergence of proposals is also very satisfactory (above 98% after about half of the recognition targets are processed). The application of the planning algorithm to training of neural networks reveals good learning results (about 50% less teacher data necessary for training in the case of two map drawings used in experiments).

1 Introduction

With the fast progressing of CAD/CAM(Computer Aided Computing / Manufacturing) system, there has been an increasing demand for automated data capture of drawings, such as maps and engineering drawings. Many systems have been successfully developed for this purpose[1],[2]. Among many others, AI-MUDAMS[3]is a system developed by the author's group to retrieve engineering drawings. It has been successfully expanded to practical systems. Since fully automated recognition for drawings with complex noise is still very difficult and manual correction of wrong recognition result takes tremendous efforts, we also developed an interactive recognition system.[4] Because the system makes use of fixed parameters, the interface result can not be reflected in interactive recognition process.

In this paper, we propose a new interactive recognition system with learning ability, which helps to improve the success rate of proposals. This system makes use of the interface results to generate proposals for new unprocessed objects. A planning algorithm is also proposed to realize fast and stable convergence of success rate. The planning algorithm is also applied to neural networks for training and proved to be effective.

2 Improved Interactive Map Drawing Recognition System

In interactive recognition system, graphical objects with little noise are processed by automated modules which utilize strictly conditioned production rules. The remaining ones are processed with rules of more flexible conditions. Proposals of recognition result are generated and presented to operator for confirmation. When a proposal is successful, the operator only need to press a "yes" button. Manual correction is necessary only when all the proposals fail. Experimental results show that costs of manual correction are greatly reduced by this approach.

The interactive recognition system previously proposed by author's group makes use of fixed parameters for generating proposals of recognition result. As a result, while being more effective than conventional systems, proposals of recognition are always generated in the same order; at the same time, the system has to present every objects to operator for confirmation. The success rate of proposals is about 70% for typical map drawings.

To improve the success rate of proposal, our new system makes use of the result of interface, that is previously processed objects. When an object is correctly recognized by interactive module, it is taken as an example and stored in example space. For unprocessed objects, their corresponding recognition result is predicted by their Euclidean distance to the existing example space. Since the more objects are processed the example space will cover more possible cases, the correct rate of proposals generated by the example space will become higher too.

Both nearest neighbor[5] and k-nearest neighbor algorithms are applied to generation of recognition proposals. K-nearest neighbor shows better performance.

Assume that previously processed objects are sorted in the example space $\{E^{(i)}\}$ (i is the class nubmer); $\{g_k\}$ is the kth feature of unknow object e, $\{f_{jk}^{(i)}\}$ the kth feature of jth example in class i. The distance between unkonw object e and class i is calculated as follows:

$$Distance(e, E^{(i)}) = Min_j\{\sqrt{\sum_k (g_k - f_{jk}^{(i)})}\} \quad (1)$$

By nearest-neighbor algorithm, the recognition proposal of an unknown object is given by finding the example space i that has the minimum distance (calculated by eqation1) from object e.

By k-nearest neighbor algorithm, first distance from the unkown object to every existing example is calculated. Then 6(determined by experiments) nearest existing examples are chosen and the number of examples that belongs to every class is counted. The class that has the maximum number of existing examples among the 6 nearest neighbors is taken as the recognition result of the unknow object.

3 Planning of Example Space and Its Optimization

Since among the graphical objects to be processed, some are similar to each other and some are very different, the success rate generated by an example space evidently depends on the contents of the example space. In other words, the correct rate of proposals depends on the order in which objects are processed. Our experimental results show that the correct rate of proposals fluctuates by 3-6% even when most of the objects are processed.

In order to realize a stable and therefore more effective learning process, we propose a planning algorithm by which the object that is most different from the obtained example space is always chosen for processing. As a result, remaining objects becomes more and more similar to the obtained example space and the success rate increases steadily. In this way, the interactive process can be switched to automated one when the success rate reaches a satisfactory level, saving the effort to process the remaining objects interactively.

Similarly, the example space needs not to store all processed objects. It can be optimized by only storing objects different from each other, thus ensuring a converged size. In this system, we use a distance threshold to decide the insertion of an example to the example space. When a proposal is incorrect, the object will be inserted into the example space unconditionally. When a proposal is correct and the distance of the processed object to example space is smaller than the threshold, it will not be inserted into example space.

4 Application of Planning Algorithm to Training of Neural Network

Neural networks have been succesfully applied in recognition of technical drawings[6][7]. After having processed certain amount of data, the obtained example space can be used to train a neural network. Since the problem with neural network is its time consuming training process, especially the process of determining the best configuration of network, we propose to apply the result of planning algorithm stated in section 3 to teacher data. For determination of configuration of networks, only objects that are most different from each other are used for initial training. As a result, the training effort can be reduced.

First, the teacher data set T (normalized to [0,1]) is sorted in the following steps:

1. Initialize the result space R as $\{\}$.
2. Find the data t that has the furthest distance from R.
3. Append t to R
4. If T is not equal to $\{\}$, repeat step 2
5. End of sorting.

Next, choose objects that have smaller distance than 1 as the selected teacher data set which is then used for determining the best configuration of the neural network.

Finally use the full set of teacher data to continue training the network determined by the selected teacher data.

In this way, the effort needed to determine the optimal configuration is reduced by the ratio of number of selected teacher data to that of total data.

5 Experimental Results

We have implemented the above mentioned system on Sun Sparc workstation in C++ language. The user interface is implemented on X-window environment. The total size of programs is about 4 thousand steps.

One typical vectorized map drawing used in experiments is shown in Fig.1. Recognition of line vector groups generated by overlapping of symbols and road lines is chosen as target of interactive operation. Some typical ones are shown in Fig.2. These vector groups have very complex variety, so that it is difficult to write explicit recognition rules for automated processing. A total of 21 parameters are used for representing the objects.

The screen layout of user interface system is shown in Fig.3. The object to be recognized is displayed at the two largest windows in the center, with different scales. The proposals are displayed at the bottom of the main window.

Fig.4 shows the learning curves of the interactive recognition process. The axis of abscissas is the number of objects processed, the axis of ordinates is the success rate of proposals. The average success rate is about 90%. The curve in dotted line show the success rate without planning. The curve in solid line shows the result of planning. The planned result show a steady increase of success rate, compared to the fluctuated one by random selection. After about half of the objects are processed, the success rate is always above 98%. (compared to 94% without planning, in dash line) Fig.5 is the result of another map with more similar objects. It shows that only 400 (out of 2500) objects need to be processed before switching to automated process, which means that only less than one sixth of objects are required to be processed before the system can switch to automated recognition with the obtained example space. Whereas for random selection, the success rate always fluctuates between 98% and 100% even when most of the objects are processed.

For the training of neural network, a back propagation network with 21 input, 1 hidden layer of 18 neurons and 4 output is used. The typical training time is about 30 minutes for every 1000 cycles of learning with a set of 1500 data. Fig.6 shows the result of neural network training with planned teacher data. By the proposed planning algorithm stated in 3, 683 are selected out of 1482 examples as initial teacher data. The axis of abscissas is the number of training cycle, the axis of ordinates the recognition rate of the trained networks. The curve in solid line is the learning curve of selected teacher data, the curve in dotted line is that of the rest. Since the recognition rate for rest of data is always higher than or equal to selected teacher data, the selected teacher data is proved to be good enough to represent the whole example space and effective for training neural network. Therefore, the effort of determining the configuration of nextwork is reduced by about 50%.

6 Conclusions

We have proposed a new type of interactive recognition system for map drawings. By making use of processed objects to generate recognition proposals of unknow objects, the efficiency of the interactive system is greatly improved when compared with that of the existing system proposed by authors' group before. By planning the order in which objects are processed, success rate of proposals can converge steadily to 100%, so that in most of the cases, only part of the objects need to be processed interactively. The planning algorithm can also be applied to data selection of teacher data for training of neural networks, which can reduce time spent for determination of optimal configuration of networks.

References

[1] R. Kasturi and L.O. Gorman: Document Image Analysis: A Bibliography, Machine Vision and Application, pp231-243, 1992

[2] Karl Tombre: Technical Drwing Recognition and Understanding: From Pixels to Semantics, Proc. of IAPR Workshop on Machine Vision Applications, pp.393-402, 1992

[3] W. Lu, Y. Ohsawa, M. Sakauchi:A Database Capture System for Mechanical Drawings Using an Efficient Multi-dimensional Graphical Data Structure Proc. of ICPR, Vol.1, pp.266-269, 1988

[4] Y. Ohsawa, Y. Takishima, M. Sakauchi, An Efficient Map Data Conversion System Using Deviced Combination of Full Automatic and Human-Assisted Recognition Phases, Trans. of IEICE, Vol. J72-DII, No.4, p545-554, 1989

[5] T. M. Cover and P.E. Hart: Nearest Neighbor Pattern Classification. IEEE Trans. on Information Theory, Vol. IT-13, no. 1, Jan. 1967.

[6] H. Arai, S. Abe, M. Nagura: Intelligent Interactive Map Recognition Using Neural Networks, 3rd ICDAR, pp.922-925, 1993

[7] T. Cheng, J.Hkan and D.Y.Y. Yun: A Symbol Recognition System, Proc. of 2nd Conf. on Document Anal. and Recog. pp.918-921, 1992

Figure 1: An Example of Vectorized Map Drawing

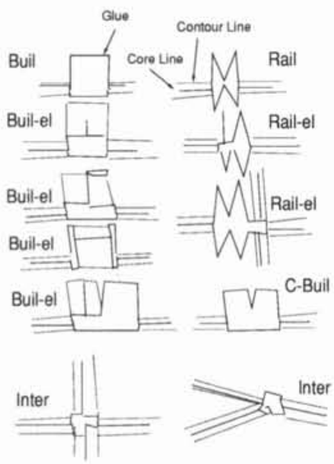

Figure 2: Examples of Glue Vectors for Recognition

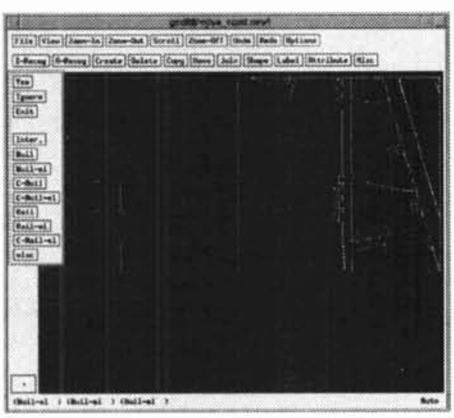

Figure 3: Layout of User Interface

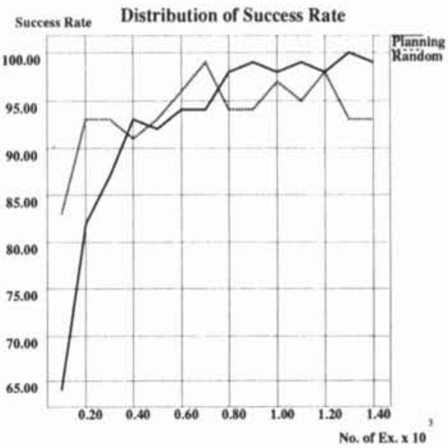

Figure 4: Learning Curves During User Interface

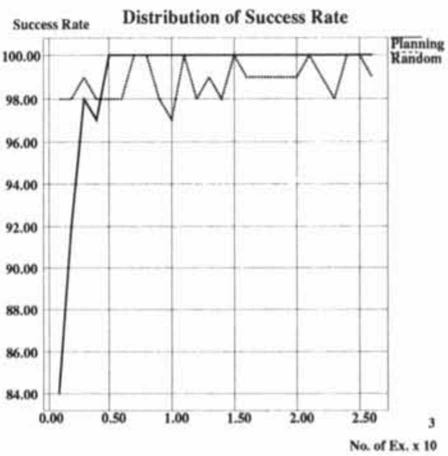

Figure 5: Another Learning Result

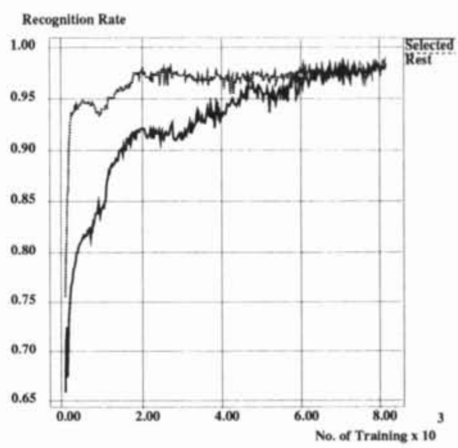

Figure 6: Training Curve of NN with Planned Teacher Data